Identification of Simple Sequence Repeats Related to Dormancy   
Induction in *Vaccinium corymbosum*

Daniel Tuerff

Laboratory Methods in Genomics

Spring 2013

**Introduction**

*Vaccinium corymbosum*, also known as the Northern highbush blueberry, is a perennial flowering plant commonly grown for its sweet tasting berries. The highbush blueberry is the most commercially produced species of blueberry in North America. More than 50 cultivars have been produced for seasonality and fruit characteristics (USDA Plants, 2012).

In order to grow blueberries more efficiently, it is important to identify the mechanisms behind plant dormancy. It has been found that the peak of dormancy is the best time for pruning plants (Bevacqua *et al*., 2012). Dormancy is also important for storing and preserving blueberries for extended periods of time during transport. Through selective breeding, it could be possible to select cultivars that enter or exit dormancy at specific times and under specific environmental conditions.

During dormancy, plants enter into a state of suspended growth with an absence of metabolic activity. Environmental factors such as lack of water or extreme temperature change can induce this state until conditions are favorable again. In addition to environmental factors, dormancy is also affected by endogenous factors such as plant hormones (Ruttink *et al*., 2007).

Abscisic acid is a plant hormone that can induce dormancy in plants. While it has many functions, abscisic acid inhibits fruit ripening, inhibits cell growth/seed germination, and down regulates enzymes needed for photosynthesis. The abscisic acid pathway begins with a pyrabactin resistance receptor (PYR/PYL) that inhibits 2C protein phosphatases (PP2C). PP2C is a negative regulator of plant dormancy. When uninhibited by PYR/PYL, PP2C removes a phosphate group from SNF1-related kinase 2 (SnRK). SnRK2 phosphorylates abscisic acid responsive element binding proteins (ABFs). Abscisic Acid Insensitive 3 (ABI3) is highly conserved amongst plants and is responsible for controlling bud growth as a transcription factor (Graeber *et al*., 2012)

In addition to genes in the abscisic acid pathway, I looked at genes causing evergrowing mutants (*evg*) in *Prunus persica*. Evergrowing mutants have a genomic deletion that affects MADS-box genes. The evergrowing deletion prevents dormancy from occurring and can be detrimental to the plant. Plants that do not enter dormancy under stressful conditions will die. Dormancy Associated MADS-box (DAM) was found to be responsible for controlling bud dormancy through regulation of gene expression (Leida *et al*., 2012). DAM6 is specifically related to peach-dormancy. Real-time RT-PCR showed reduced DAM6 transcript levels during a rise in growth competence of flower buds (Leida *et al*., 2012). DAM6 could be used as a marker for chilling requirements and dormancy timing. It would be possible for blueberry growers to know which crops

My project consisted of BLASTing genes related to dormancy against 454 blueberry genome scaffolds. Through the BLAST results, I was able to find simple sequence repeats (SSRs). The BLAST results allowed me to generate primers using the *Vaccinium* website. These primers could then be used by blueberry cultivators for selective breeding. During the project, I also tested the GenSAS website to see if it would be useful for our blueberry project. I was assigned to look at the tRNA scan function of the website.

**Materials and Methods**

I began the assignment with the search for articles on dormancy induction in eudicots. I searched for organisms that were similar to *Vaccinium corymbosum* with National Center for Biotechnology Information (NCBI) and Google. *Arabidopsis thaliana* and *Prunus persica* were commonly used because of the extensive amount of research in dormancy. Genes that were related to the abscisic pathway, metabolism, oxidation-reduction, and signaling/transcription in plants were picked from these papers on dormancy. I also found genes involved in signaling pathways through the Kyoto Encyclopedia of Genes and Genomes (KEGG). After finding specific genes and proteins, I found the amino acid sequences corresponding with the proteins using the protein search in NCBI. I omitted proteins that did not have a full sequence.

**BLAST**

The next step required using Apple’s Terminal app to BLAST the amino acid sequence against the blueberry scaffolds. To do this, I used the command “cd” (file location here)” without the parenthesis to allow the program to locate the folder containing the 454 scaffolds. Next, I used the command “./bin/makeblastdb -in (scaffold txt file name here) -input\_type fasta -dbtype nucl -title blueberry\_Genome” to create a database from the 454 scaffolds. This loaded the previously gathered 454 blueberry scaffolds for BLASTing against my dormancy genes.

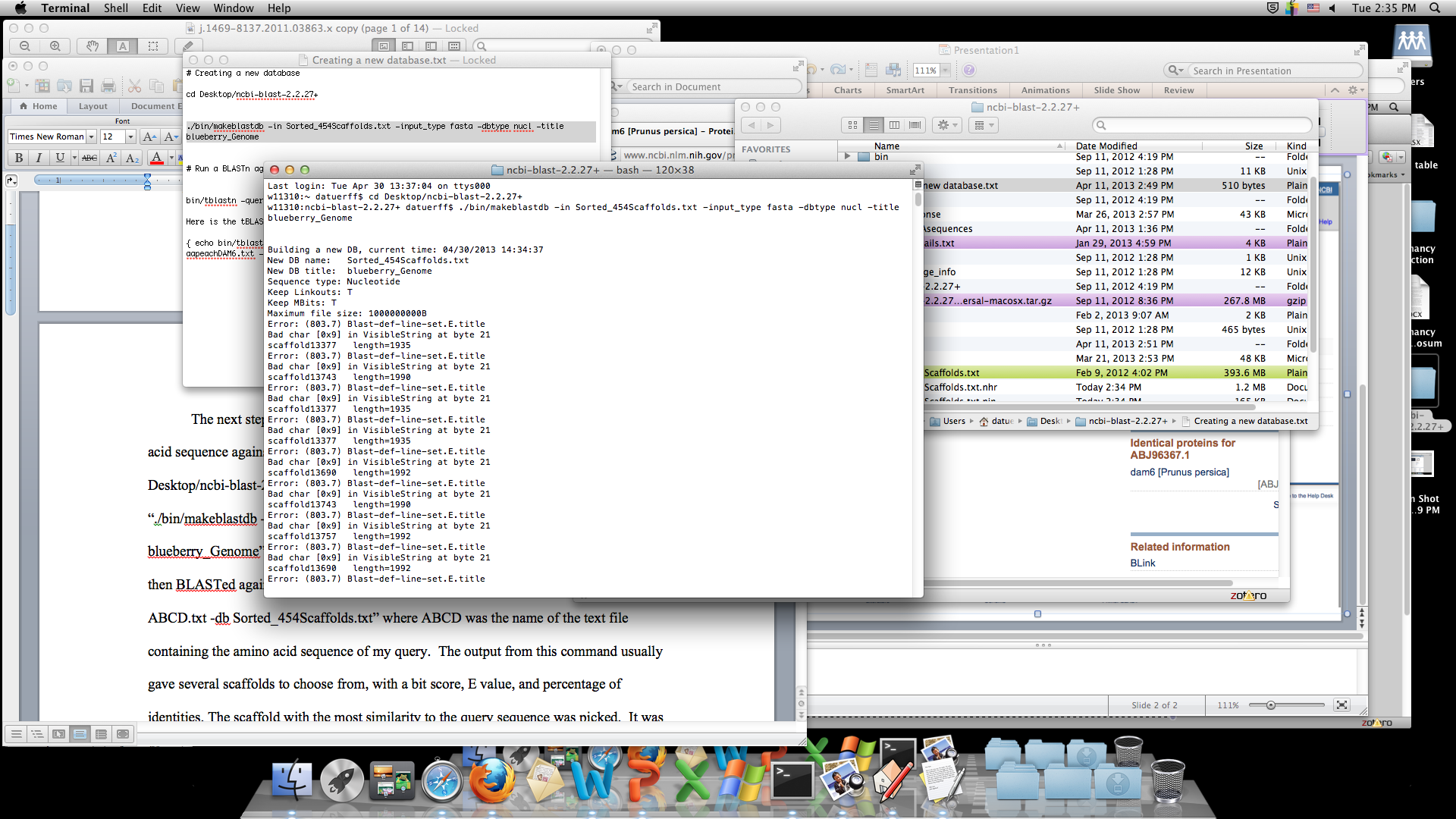


Figure 1. Terminal window showing commands used to create a database for BLAST.

I BLASTed the amino acid sequences against the new database using the command “bin/tblastn -query (Amino acid sequence in FASTA format) -db Sorted\_454Scaffolds.txt.” The command uses tBLASTn to compare the protein query amino acid sequence against the nucleotide sequence from the scaffolds. The output from this command usually gave several scaffolds to choose from, with a bit score, E value, and percentage of identities. I picked the scaffold with the lowest E value because it had the most similarity to my query. It was important to note which nucleotide the query sequence started on the scaffold (Figure 3).

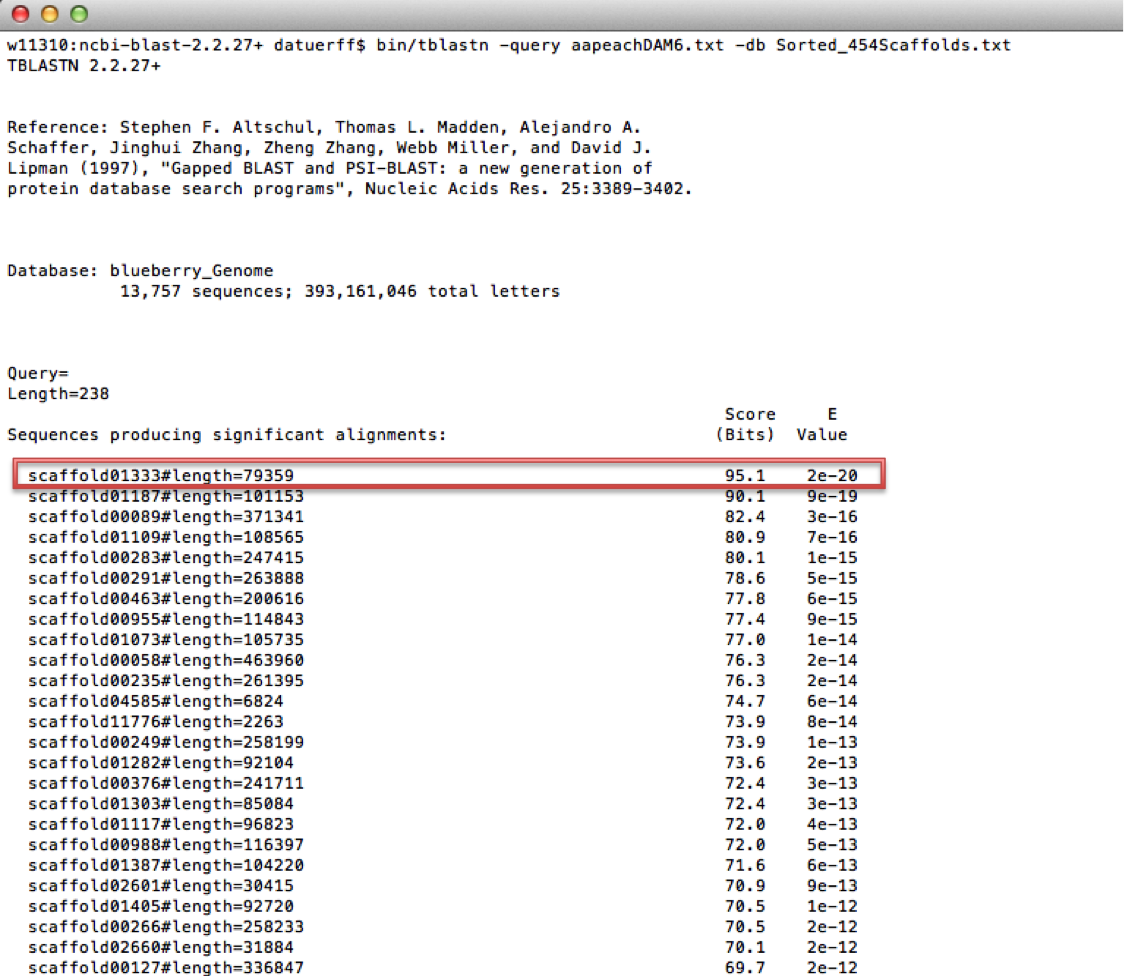


Figure 2. tBLASTn results. Scaffold with lowest E Value is highlighted in red.

The screenshot below shows the starting location of my query sequence on the scaffold. In this case, my query sequence starts at the 18535th nucleotide on scaffold 1333.

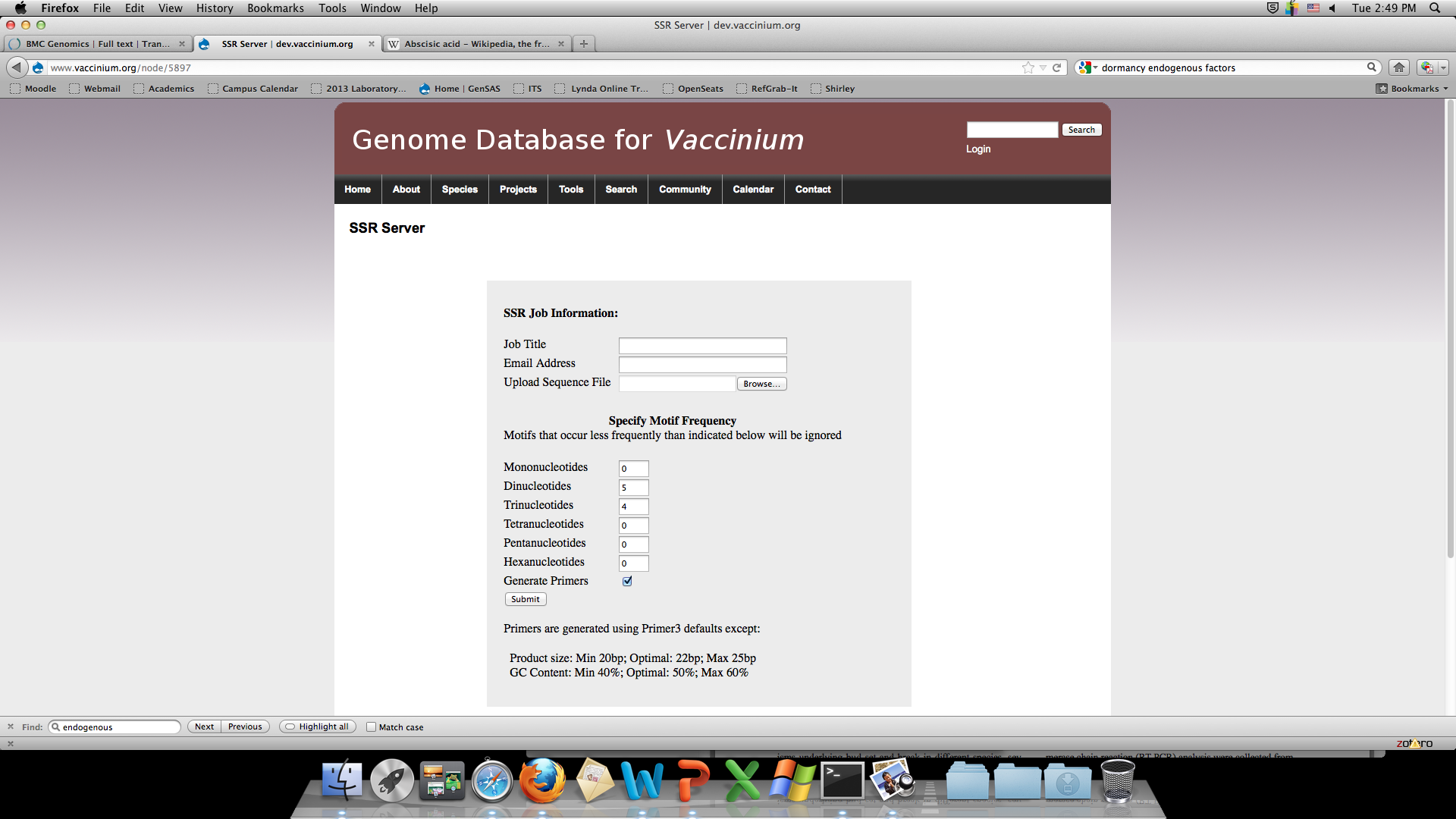


Figure 3. tBLASTn results continued. The query start location is highlighted in red.

***Vaccinium* Database Search**

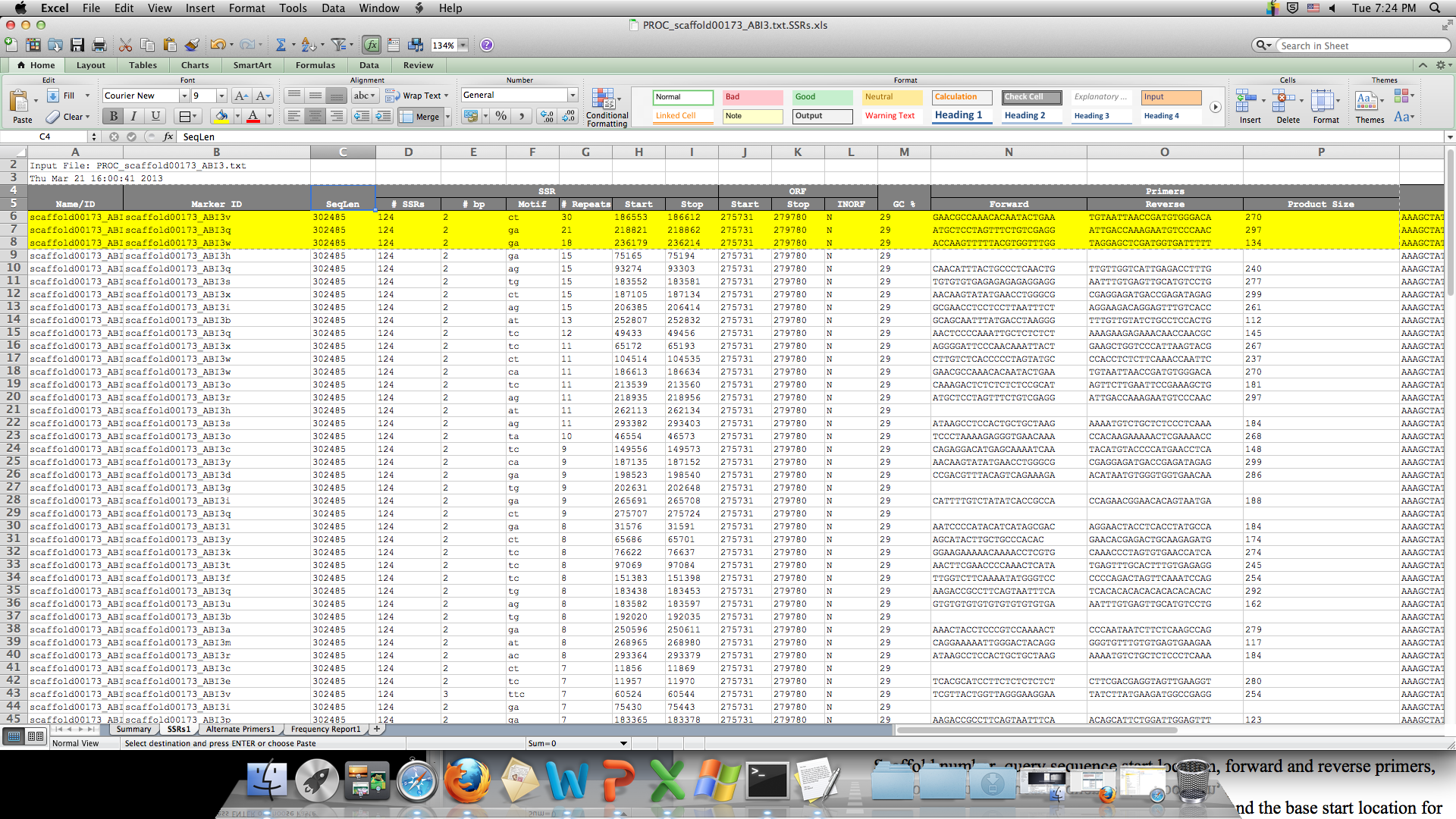
Using the scaffold with the lowest E-value, I searched the provided 454 blueberry scaffolds for the scaffold containing the gene of interest. I used the “ctrl + F” function to bring up a search window in the .txt file containing the scaffolds. I typed in “scaffold(number here)” in the search box to search for the scaffold identified by the tBLASTn. The entire scaffold was copied into a .txt file to be uploaded on the *Vaccinium* website. To upload my scaffold, I went to the *Vaccinium* website, over the tools tab and clicked SSR. Figure 4 shows the SSR Server that The sequence file was uploaded with values of 0 for tetranucelotides, pentanucleotides, and hexanucelotides.

Figure 4. Vaccinium Genome Database SSR search (Genome Database for Vaccinium. Tetra through Hexanucleotides (highlighted in red) should have values of zero.



**Primers**

After submitting a job, a link was sent back to me through an email. The link took me to a webpage with multiple other links. I downloaded the link with the excel file. The excel file was sorted by number of repeats so that the greatest number of repeats was toward the top. Primer pairs were picked by both number of repeats and proximity to the query sequence. My first priority was to pick the primers that had more than 15 repeats. If there were none, I tried to find primers that were at least within 10,000 bases of my gene. Using these guidelines, I was able to generate three sets of primers.



Repeats

Primers

Figure 5. Excel spreadsheet showing SSR search results. Repeats and primers are located with red arrows.

Scaffold number, query sequence start location, forward and reverse primers, nucleotide repeats, number of repeats, PCR product size, and the base start location for each SSR were recorded into an organized word file.

Example of a finished report:

Found in Scaffold 00173 (query sequence starts at base 74854 on scaffold)

1)

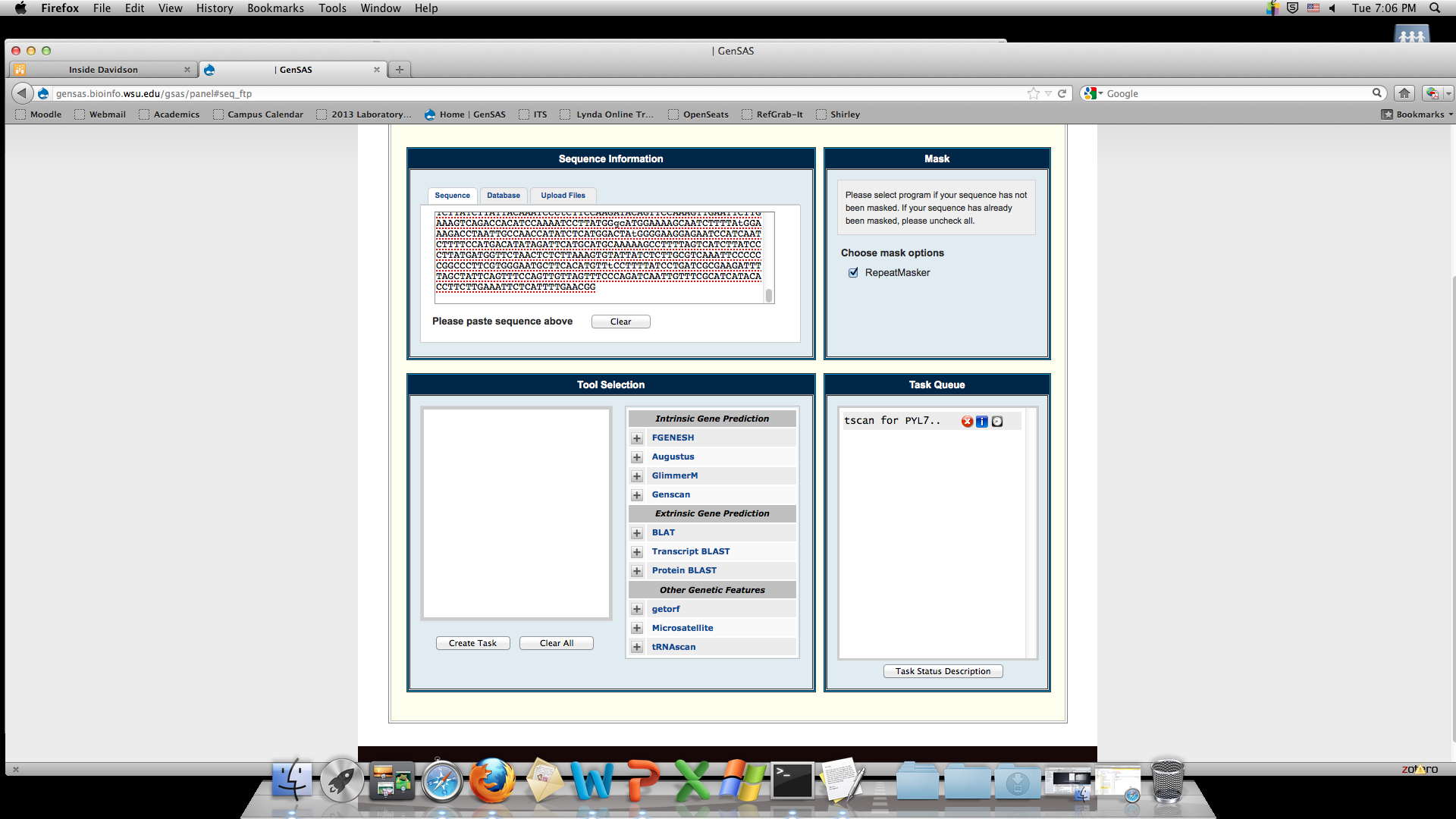
For Primer GAACGCCAAACACAATACTGAA

Rev Primer TGTAATTAACCGATGTGGGACA

repeats (ga) x30 PCR product =270 bp & start at base 186553

**GenSAS**

To test the GenSAS website, I first logged on and clicked “Use GenSAS” (GenSAS, 2013). I copied the scaffold sequence identified for a dormancy related gene into the sequence box. I then added tRNAscan as a task, hit “create task,” and hit the feather icon in the task queue to start the task. The feather turned blue to let me know that the task had started. I was sent an email telling me my scan was complete, and logged in to look at the results.Figure 6. GenSAS tRNAscan. The tRNAscan tool is framed in red.



**Results**

23 total SSRs were obtained. For each SSR, three sets of forward and reverse primers were generated. The genes were picked from the abscisic pathway, metabolism, oxidation-reduction, and dormancy-related signaling/transcription in plants. The majority of organisms were either *Arabidopsis thaliana* or *Prunus persica.*

|  |  |
| --- | --- |
| Gene | Organism |
| ABI3 | *Arabidopsis thaliana* |
| FUS3 | *Arabidopsis thaliana* |
| LEC1 | *Arabidopsis thaliana* |
| LEC2 | *Arabidopsis thaliana* |
| PP2C | *Arabidopsis thaliana* |
| PYL/PYR | *Arabidopsis thaliana* |
| SNRK2 | *Arabidopsis thaliana* |
| Cytochrome p450 | *Populus trichocarpa* |
| DAM5 | *Prunus persica* |
| DAM6 | *Prunus persica* |
| NAC domain protein IPR003441 | *Populus trichocarpa* |
| Zinc finger protein | *Camellia sinensis* |
| GRAS family transcription factor | *Populus trichocarpa* |
| NAC domain protein NAC1 | *Gossypium hirsutum* |
| Mitogen activated protein kinase kinase kinase | *Ricinus communis* |
| Transcription factor AP2-EREBP | *Lotus japonicus* |
| DAM4 | *Prunus persica* |
| FLC | *Arabidopsis thaliana* |
| UDP-galactose 4 epimerase | *Cyamopsis tetragonoloba* |
| S-like ribonuclease | *Prunus dulcis* |
| Acyl:coa ligase | *Populus trichocarpa* |
| Strictosidine synthase family protein | *Brassica napus* |
| chs-like protein | *Populus trichocarpa* |
| Carboxyl-terminal protease | *Zea mays* |
| Chalcone synthase family protein | *Arabidopsis halleri subsp. Gemmifera* |
| Xyloglucan endotransglucosylase/ hydrolase 5 | *Malus x domestica* |

Table 1. SSRs identified for each gene except for FLC (shown in red). Genes were obtained from several different organisms.

An SSR for the transcription factor Flowering Locus C (FLC) was not obtained because the scaffold was too small. I found multiple SSRs for genes involved in hormone regulation, maturation regulation, signaling, transcription, and metabolism related to dormancy.

Unfortunately, GenSAS was unable to find any tRNAs in my scan. It is possible that this tool works under certain conditions. However, my tests did not yield any results. The website is currently unable to run any tasks I give it (5/7/2013).

**Discussion**

I was able to obtain SSRs for 23 different dormancy-related genes. The majority of the genes were related to either the abscisic acid pathway or evergrowing mutants. I was able to obtain a scaffold for FLC, however no SSR primers were generated from the *Vaccinium* website. Future research may be able to obtain primers for FLC if a larger scaffold is generated through blueberry genome sequencing. FLC is an important gene because responsible for regulating the developmental life cycle in *Arabidopsis thaliana*.

SSRs that I successfully generated from this research could be used to identify plants for crossbreeding by tracking different alleles. Primers generated from my research could be used to create PCR product from a young blueberry plant. If PCR product size is different for a certain allele, it could be used to screen plants for specific genotypes before they are fully grown. This would speed up the screening process to get desired phenotypes such as increased or decreased time of dormancy. My research will help blueberry growers create cultivars with desired traits. With increased control of blueberry dormancy timing, growers can accurately know when their blueberries will grow. Future research should continue the search for dormancy related genes in *Vaccinium corymbosum*.

**References**

Bevacqua D, Génard M, Lescourret F. A simple model to predict the probability of a peach (Prunus persicae) tree bud to develop as a long or short shoot as a consequence of winter pruning intensity and previous year growth. PLoS ONE. 2012;7(12):e52185.

Genome Database for Vaccinium. Plants for Human Health Institute. 07 May 2013. <http://www.vaccinium.org/>.

GenSAS. 07 May 2013. <http://gensas.bioinfo.wsu.edu/>.

Graeber K, Nakabayashi K, Miatton E, Leubner-metzger G, Soppe WJ. Molecular mechanisms of seed dormancy. Plant Cell Environ. 2012;35(10):1769-86.

KEGG: Kyoto Encyclopedia of Genes and Genomes. 07 May 2013. <http://www.genome.jp/kegg/>.

"Highbush Blueberry." *USDA PLANTS*. N.p., n.d. Web. 07 May 2013. <http://plants.usda.gov/java/>.

Leida C, Conesa A, Llácer G, Badenes ML, Ríos G. Histone modifications and expression of DAM6 gene in peach are modulated during bud dormancy release in a cultivar-dependent manner. New Phytol. 2012;193(1):67-80.

Leida C, Terol J, Martí G, et al. Identification of genes associated with bud dormancy release in Prunus persica by suppression subtractive hybridization. Tree Physiol. 2010;30(5):655-66.

National Center for Biotechnology Information. U.S. National Library of Medicine. 07 May 2013 <http://www.ncbi.nlm.nih.gov/>.

Ruttink T, Arend M, Morreel K, et al. A molecular timetable for apical bud formation and dormancy induction in poplar. Plant Cell. 2007;19(8):2370-90.

**Acknowledgments**

Thank you Dr. Allan Brown at NCSU, Dr. Jeannie Rowland with the USDA, Dr. Doreen Main at WSU, Dr. Malcolm Campbell at Davidson College, and the BIO343 class.

**Appendix**

ABI3 Found in Scaffold 173 (query sequence starts at base 74854 on scaffold)

1)

For Primer GAACGCCAAACACAATACTGAA

Rev Primer TGTAATTAACCGATGTGGGACA

repeats (ga) x30 PCR product =270 bp & start at base 186553

2)

For Primer ATGCTCCTAGTTTCTGTCGAGG

Rev Primer ATTGACCAAAGAATGTCCCAAC

repeats (ga) x21 PCR product =297 bp & start at base 218821

3)

For Primer ACCAAGTTTTTACGTGGTTTGG

Rev Primer TAGGAGCTCGATGGTGATTTTT

Repeats (ga) x18 PCR product =134 bp & start at base 236179

Acylcoa Found in Scaffold 1173 (query sequence starts at base 36878 on scaffold)

1)

For Primer ATTCAACGGCTCAGATTCACTT

Rev Primer TACAAATCCAAACACACGGAAC

repeats (ta) x29 PCR product =251 bp & start at base 87799

2)

For Primer TGTGAAATGGAGGTACCCTTCT

Rev Primer CAGATGAATTGCTGAAACCTGA

repeats (ct) x15 PCR product =187 bp & start at base 76484

3)

For Primer CATTGGGCTGACATTAAGGTTT

Rev Primer TCTGCATGTTTGTTGGGTTTAG

Repeats (tg) x10 PCR product =210 bp & start at base 6200

Carboxyl terminal processing protease Found in Scaffold 2378 (query sequence starts at base 10045 on scaffold)

1)

For Primer AAGGCTCAAGACAACAATCCAT

Rev Primer GCTCCATATACGAAAAGCTTGG

repeats (ag) x10 PCR product =102 bp & start at base 18932

2)

For Primer TTTTTATGTATCCCCGGAACC

Rev Primer CGAAATGCTGAAACTTGTCTTG

Repeats (ac) x7 PCR product =186 bp & start at base 60

3)

For Primer TGCTGGAGAAAAAGAAACACTG

Rev Primer ATGATTGGAGCTCTGATACCGT

repeats (ct) x6 PCR product =213 bp & start at base 2896

Chalcone synthase family protein Found in Scaffold 491 (query sequence starts at base 132065 on scaffold)

1)

For Primer GCCAACTATAAGCCTGAACCAC

Rev Primer ATAGCCTGGGTTTCACAGATTG

repeats (ag) x16 PCR product =281 bp & start at base 92449

2)

For Primer GAATAGCAGGGTCTTTCCAATG

Rev Primer ATTTCCGGACACTATTACCCCT

repeats (ag) x14 PCR product =239 bp & start at base 152527

3)

For Primer GCCAACTATAAGCCTGAACCAC

Rev Primer ATAGCCTGGGTTTCACAGATTG

Repeats (tg) x12 PCR product =281 bp & start at base 92425

Chs-like protein Scaffold 814 (query sequence starts at base 40156 on scaffold)

1)

For Primer AGGTCTTTGTTTCCAACCTTGA

Rev Primer CGGTTTGGTTTGGTTCATAAGT

repeats (ga) x17 PCR product =232 bp & start at base 13893

2)

For Primer GAATTTGCAGGAATTATGGCTC

Rev Primer GTTCCCCATTCTTCTCCTCTCT

repeats (ct) x11 PCR product =170 bp & start at base 28175

3)

For Primer AGACAACCTTCCCAAACTCAAA

Rev Primer TGCAGTATGCTTTCTGACCACT

Repeats (ca) x10 PCR product =279 bp & start at base 10144

Cytochrome p450 Found in Scaffold 586 (query sequence starts at base 92166 on scaffold)

1)

For Primer ATAGCGGCTCTCCAAAAACAT

Rev Primer ATCTAGGGTTTCGAGTCCACCT

repeats (ag) x21 PCR product =234 bp & start at base 136135

2)

For Primer AGCGTATTCCATTCTGGACTGT

Rev Primer ATATTTGGCACCTCAATTGCTC

repeats (ta) x15 PCR product =293 bp & start at base 56735

3)

For Primer GTACGAGAAAAACTTGGATGCC

Rev Primer TATCCAGAGATTCCAGACCCAT

Repeats (ga) x9 PCR product =267 bp & start at base 50281

Cytochrome p450 Found in Scaffold 586 (query sequence starts at base 92166 on scaffold)

1)

For Primer ATAGCGGCTCTCCAAAAACAT

Rev Primer ATCTAGGGTTTCGAGTCCACCT

repeats (ag) x21 PCR product =234 bp & start at base 136135

2)

For Primer AGCGTATTCCATTCTGGACTGT

Rev Primer ATATTTGGCACCTCAATTGCTC

repeats (ta) x15 PCR product =293 bp & start at base 56735

3)

For Primer GTACGAGAAAAACTTGGATGCC

Rev Primer TATCCAGAGATTCCAGACCCAT

Repeats (ga) x9 PCR product =267 bp & start at base 50281

DAM4

Found in Scaffold 01333 (query sequence starts at base 18544 on scaffold)

1)

For Primer AGCCGTAGGAGAACAAGATCAA

Rev Primer GCACTCCTAGGTCCGAAGTATC

repeats (ac) x10 PCR product =242 bp & start at base 73805

2)

For Primer ATCATGCACCGCCCTATTATAC

Rev Primer GAGGCACTTTTATATGGGTTGG

repeats (tc) x7 PCR product =244 bp & start at base 27010

3)

For Primer TATGGTCAACAACCGCTACATC

Rev Primer TTTATTAGGGCTTGGTCTCGAA

Repeats (gca) x9 PCR product =210 bp & start at base 30618

DAM5 Found in Scaffold 1187 (query sequence starts at base 47701 on scaffold)

1)

For Primer GCTACTGTTCTCGTCTCCTCGT

Rev Primer TCAATAGAACCTTCCACCGACT

repeats (ct) x13 PCR product =116 bp & start at base 78977

2)

For Primer AATGTTTTGTGTCTCCCAATCC

Rev Primer GGCACAACGCTAGTTGATAAAG

repeats (ga) x12 PCR product =275 bp & start at base 95120

3)

For Primer CAAATGGTGATAGCAAACATGG

Rev Primer CTCAAAGAAAATTCCCAAGACG

Repeats (tta) x7 PCR product =274 bp & start at base 57880

DAM6 Found in Scaffold 01333 (query sequence starts at base 18535 on scaffold)

1)

For Primer TGGGTAGAATCAGAGAAAACGG

Rev Primer AAAATGACGGGTGGTCCTATTT

repeats (ag) x18 PCR product =284 bp & start at base 58236

2)

For Primer AGCCGTAGGAGAACAAGATCAA

Rev Primer GCACTCCTAGGTCCGAAGTATC

repeats (ac) x10 PCR product = 242 bp & start at base 73824

3)

For Primer TATGGTCAACAACCGCTACATC

Rev Primer TTTATTAGGGCTTGGTCTCGAA

Repeats (gca) x9 PCR product =210 bp & start at base 30592

FUS3LEC2 Found in Scaffold 00219 (query sequence starts at base 74050 on scaffold)

1)

For Primer ACTGATGGAGGAGCACGATTAT

Rev Primer TTTTTGAGGTGCAATGTGACTC

repeats (tc) x10 PCR product =262 bp & start at base 163071

2)

For Primer TGGGTTCCATTACTACCTAGCG

Rev Primer CAAATTGAAACGGAAGGAGTGT

repeats (ct) x PCR product = 257 bp & start at base 27318

3)

For Primer GGGGTACTGGGTTCCATTACTA

Rev Primer TAGGTTTGAAAATCGAGAGCGT

Repeats (tc) x PCR product =285 bp & start at base 22139

GRAS family protein Found in Scaffold 567 (query sequence starts at base 24082 on scaffold)

1)

For Primer AAACAAGCATCAACGCAGAGTA

Rev Primer AAGCTTGTGGGGTCATAGGATA

repeats (ag) x13 PCR product =219 bp & start at base 29639

2)

For Primer ACGTACCCTCTCAACGAGTCAT

Rev Primer ATCAAAAATCCCTTACTGGGGT

repeats (tc) x12 PCR product =216 bp & start at base 23439

3)

For Primer GATTCCCTATCCAGCTTACGTG

Rev Primer TTCCACTAGTTTAATGTCCGGC

Repeats (ag) x6 PCR product =118 bp & start at base 20805

LEC1 found in scaffold 00036 (query sequence starts at base 301135 on scaffold)

1)

For Primer GCACACAAAGGAGTTTCACTCA

Rev Primer GTTTCGTTTCCTTTCGTTTCCT

repeats (ag) x30 PCR product =187 bp & start at base 95860

2)

For Primer ACTCGTTGCGAATAAATCCACT

Rev Primer CACCCCAATTCTCCAAAGGTA

repeats (ag) x14 PCR product =172 bp & start at base 473440

3)

For Primer TACTCAGTGGACATTGCTTGCT

Rev Primer CCTCTTCCTTCTCCTCCTCTTC

Repeats (ag) x22 PCR product =245 bp & start at base 164682

MAPKKK3 Found in Scaffold 394 (query sequence starts at base 135452 on scaffold)

1)

For Primer TTGAGTGCATTTCTCTCTGGAA

Rev Primer TAATGATGACATGGATTCACCG

repeats (ga) x13 PCR product =125 bp & start at base 170993

2)

For Primer CATTAACCCTACCCCATCTTCA

Rev Primer GAAATAAAACCCGTTACGGACA

repeats (ct) x12 PCR product =235 bp & start at base 28356

3)

For Primer ACTGCTACGATCAAAACTGGGT

Rev Primer ATAATGAAGTCGGGGTTTGATG

Repeats (tc) x11 PCR product =277 bp & start at base 52896

NAC domain protein NAC1 Found in Scaffold 72 (query sequence starts at base 258035 on scaffold)

1)

For Primer TGGGTACAAAACCTTAGCTTGG

Rev Primer TCTCGTACCCAAAGAAATGACC

repeats (ct) x25 PCR product =164 bp & start at base 347939

2)

For Primer AACAATCGGCCCTTGTTAGATA

Rev Primer TATTCTAAATCCAGGGCCTTCA

repeats (tg) x19 PCR product =129 bp & start at base 194575

3)

For Primer TTCATAATTTGGGGAGGGAGTA

Rev Primer AGGCCGACTTTCTCTTCTTCTT

Repeats (gt) x12 PCR product =256 bp & start at base 142010

NAC domain protein IPR003441 Found in Scaffold 01415 (query sequence starts at base 34984 on scaffold)

1)

For Primer AGATTATTAGCGTGAAGCTCCG

Rev Primer GATGCCATCATACGAAAGGACT

repeats (ga) x19 PCR product =214 bp & start at base 18700

2)

For Primer AGGGGTCCATAGCACATAAAAA

Rev Primer ATGTCGGAGGAAGTCGAAACTA

repeats (ag) x14 PCR product =244 bp & start at base 54313

3)

For Primer TACCAAATACCAATTTGAGGCG

Rev Primer ATCTAAATACTCCCTCCGTCCC

Repeats (ta) x7 PCR product =127 bp & start at base 44338

PP2C Found in Scaffold 644 (query sequence starts at base 26276 on scaffold)

1)

For Primer CTCCCAAAAACTCTGCACCTAC

Rev Primer CCCCATTAATTTACATAGGCGA

repeats (ct) x16 PCR product =259 bp & start at base 1879

2)

For Primer GCATGCCTGATCACAATTAGAA

Rev Primer ATTTCCAGATGGTCTCTTACCG

repeats (ga) x9 PCR product =175 bp & start at base 70686

3)

For Primer ATTTCTGGTCTTTCCTTCACCA

Rev Primer AGTTCTCGAATGGAAGCAATGT

Repeats (tc) x13 PCR product =133 bp & start at base 122220

PYL7 Found in Scaffold 00766 (query sequence starts at base 30762 on scaffold)

1)

For Primer CCCCAATTCAGGTCTATCAAAA

Rev Primer CTAAGAGCATCTCCAACTGCCT

repeats (ga) x7 PCR product =297 bp & start at base 3418

2)

For Primer CTTTCTCCTTTCACCCAAACAC

Rev Primer TTGCATATCCGTCTCAAGAAGA

repeats (ct) x8 PCR product =191 bp & start at base 87340

3)

For Primer TCTGTCATCCTGCACAATATCC

Rev Primer GAAAGTTGAGTTTTCGAATGGG

Repeats (ct) x6 PCR product =257 bp & start at base 66733

S like Ribonuclease Found in Scaffold 484 (query sequence starts at base 175226 on scaffold)

1)

For Primer CAGGACCACAAATGAATTAGCA

Rev Primer AACTTACCGTAAACGACCTCCA

repeats (ag) x16 PCR product =139 bp & start at base 78786

2)

For Primer GGACTAACAATGAAGGGTGGAG

Rev Primer GGCTCAGAAAATAGCAGCACTT

repeats (ga) x11 PCR product =265 bp & start at base 85952

3)

For Primer TTACACCTCAGGATGTGCAGAC

Rev Primer CCCCCAAATTAAAGGGTAAAAC

Repeats (ga) x10 PCR product =277 bp & start at base 87363

SNRK2 found in Scaffold 00548 (query sequence starts at base 150361 on scaffold)

1)

For Primer TATGTGTGCTAACCTGCCATTC

Rev Primer CCAAAACCTCGAAATCAAGAAC

repeats (ct) x12 PCR product = 247 bp & start at base 77319

2)

For Primer GCTCCTAGAAAGAAACCTACCTCA

Rev Primer TTAAGGCACCAAGAACTCCACT

repeats (ag) x21 PCR product = 291 bp & start at base 87552

3)

For Primer TGGAATTAAGTTCGCTAACCGT

Rev Primer TTGCAGAAACCACTTAACCCTT

Repeats (ag) x PCR product = 286 bp & start at base 175818

Strictosidine synthase family protein Found in Scaffold 182 (query sequence starts at base 68497 on scaffold)

1)

For Primer CTGAAATGTACCAAAACGACCA

Rev Primer AAACAACCTGGGTAAACAATGC

repeats (ga) x16 PCR product =167 bp & start at base 297782

2)

For Primer CTAGGTCGCTCACATTCTTCAA

Rev Primer GAAATAGGAGGATGGTGCGTAG

repeats (ca) x12 PCR product =262 bp & start at base 144969

3)

For Primer AAGTACCCCAAGATTGGAAGTG

Rev Primer GCTCGCCCTTGATATATTTTTG

Repeats (ga) x12 PCR product =262 bp & start at base 235417

UDP galactose-4 epimerase Found in Scaffold 658 (query sequence starts at base 70755 on scaffold)

1)

For Primer CAGGGACTACAAATACGCTTCC

Rev Primer TCACGAAACATATACGGCAAAG

repeats (ct) x9 PCR product =272 bp & start at base 44180

2)

For Primer CAACACATGAACATTACCCACC

Rev Primer GACTTGTTGGGGGAGAGTGTAG

repeats (ag) x8 PCR product =285 bp & start at base 51379

3)

For Primer TGAGTTGCCAAGAATTGTGTTC

Rev Primer GCATTAGATGGGTTTTCTGGAG

Repeats (ga) x9 PCR product =263 bp & start at base 143206

Xyloglucan endotransglucosylase/ hydrolase 5 Found in Scaffold 1511 (query sequence starts at base 36878 on scaffold)

1)

For Primer GATGGACTTGTGTGTGTGTGTG

Rev Primer TGAGCCTTTGAAAACCTCTCTC

repeats (ag) x25 PCR product =126 bp & start at base 27538

2)

For Primer TTCTGGTCTTGGTACGAATGTG

Rev Primer CGTAATCAACAGTTCACTCCCA

repeats (ga) x12 PCR product =120 bp & start at base 33443

3)

For Primer GAGGGTGAACAAATTAAATGGC

Rev Primer TGCTTCTCTCCCAAATTTGACT

Repeats (tc) x10 PCR product =283 bp & start at base 16624

Zinc finger protein Found in Scaffold 2129 (query sequence starts at base 12347 on scaffold)

1)

For Primer ACCCTAACAAAATCAAAACCCG

Rev Primer TCCAAACACGATTTCTAACCCT

repeats (ga) x21 PCR product =152 bp & start at base 40907

2)

For Primer ACTACCTACCCCTAACCCCGTA

Rev Primer AATTAGGTGTGTTGCTGGGAAT

repeats (ct) x16 PCR product =242 bp & start at base 39654

3)

For Primer GGGAAGAGAGAGAGGAAAGGAG

Rev Primer AACCGTAGACTTTTGGTTTGGA

Repeats (ga) x8 PCR product =280 bp & start at base14865